INFERRING MULTIPLE COFFEE FLOWERINGS IN CENTRAL AMERICA USING FARMER DATA IN A PROBABILISTIC MODEL

Leonel Lara-Estrada[1], Luis Enrique Sucar[2], Livia Rasche[3]
[1]University of Greenwich, UK
[2] Instituto Nacional de Astrofísica Óptica y Electrónica, México
[3]University of Hamburg, Germany

Introduction

Climate and weather are key factors in triggering flowering events. After a period of water stress, rainfall might break the dormancy of flowering buds, leading to flowering days after. However, there is uncertainty about the occurrence and magnitude of flowering due to rainfall on a given day during the water-stress months. Therefore, in this study, we developed a Dynamic Bayesian Network model to forecast the likelihood and intensity of multiple flowering events using observed data on flowering and precipitation from coffee farms in the Pacific Region of Nicaragua, Central America.

Methods

- Pacific Central America: dry season (5-6 months, Dec-April) with the lowest rate of precipitation and high temperatures.
- Data collected by Lara-Estrada et al. (2012)
- Farms' records on flowering dates and intensity, daily rainfall were used to estimate *flowering intensity*, (flowering) *month*, *rainfall that induces flowering*, and *days to flowering*.
- Model structure: literature and expert knowledge. Four flowering events/year max. (in >90% flowering records).
- Model training: Counting-Learning Algorithm using 53 years of data (1943-1998) x 1 farm.
- Validation for *flowering intensity* and *days to flowering*: Spherical Payoff (SP) using 4-5 years of data (1999-2010) x 4 farms. SP ranges from 0 to 1, with 1 being the best fit.
- Sensitivity analysis to quantify the changes in the values for flowering intensity due to findings (changes) in other model variables.

Bayesian Networks (BNs) are multivariate statistical models that consist of two main components: a directed acyclic graph (model structure) and conditional probability tables (model parameters). Together, they compactly represent the joint probability distribution. An acyclic graph consists of a set of nodes (variables) linked by arcs; the arcs define the conditional dependencies. A BN becomes a Dynamic BN when the same phenomenon occurs in the model at different time slides. E.g., multiple flowerings.

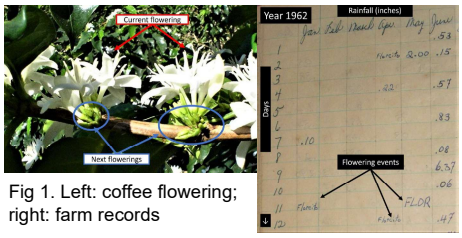

Fig 1. Left: coffee flowering; right: farm records

Results and Discussions

Sensitivity Analysis
Flowering intensity is highly influenced by the next (F1, F2) or previous (F3, F4) flowerings, then Rain IF, month, and Days IF. So, the intensity of initial flowerings influences further ones. *Day to flowering*: Month and Rain IF. Less sensitive than Flowering intensity.

Model performance (4 farms)
A good performance in estimating flowering intensity (SP=0.78), compared to *days to flowering* (SP=0.45). The last could be influenced by the uncertainty associated with Rain IF (the sum of rainy days in some cases) and the node's state breaks.

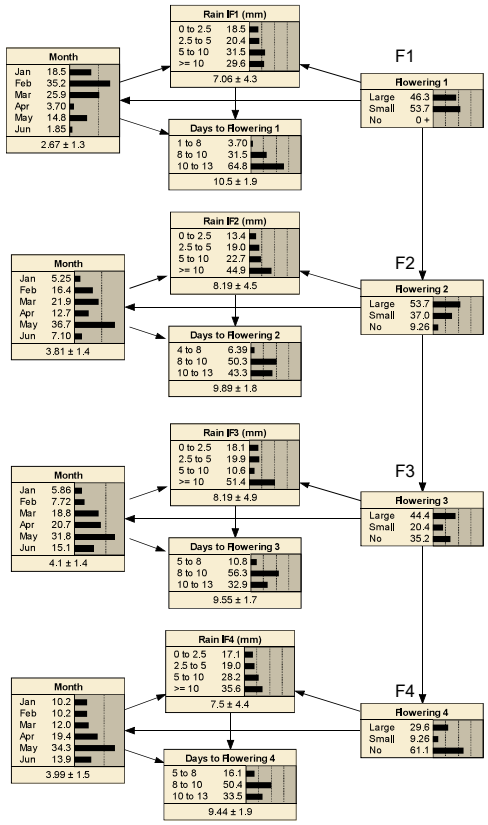

Fig 2. Dynamic Bayesian Network mode for coffee flowering, Pacific Region of Central America. The model depicts the prior probabilities for each variable prior to findings entry. If a user enters the month and rainfall data (findings), the flowering intensity and the days to flowering will be inferred.
Rain IF: rainfall that induces a flowering,
Days to flowering: days to flowering after Rain IF,
Flowering: Flowering intensity (F1-F4).

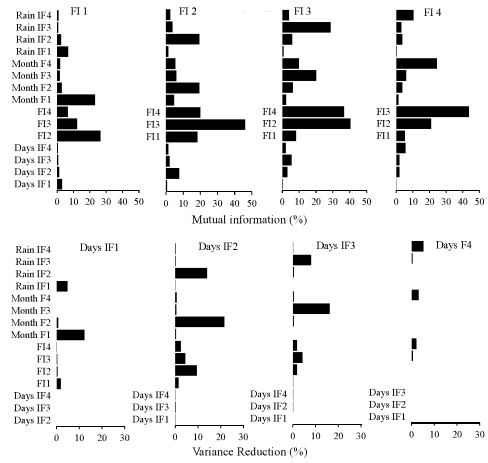

Fig 3. Sensitivity analyses results for the flowering intensity (top) and days to flowering (bottom) for the four flowering events. Rain IF: rainfall that induces a flowering, Days IF: days to flowering after Rain IF, FI: Flowering intensity

Testing and learning from the model

The model learned and infers the flowering events triggered by rain across different months, based on climate patterns observed over 5 decades of data.

Different cases and scenarios were tested, what if, and the most probable outcomes for those situations were obtained, e.g. :

- The earlier in the year the first flowerings occurred (Jan-Feb), the smaller they would be, even if rains >10 mm;
- From March onwards, rains >10 mm are most likely to produce large flowerings.
- If there are four flowerings, one or two will be large. If only two flowerings, they tend to be large.
- F3 and F4 in May will be large.
- If four flowerings have not occurred by May, it is unlikely a F4 will occur in June; but an F3 in June is possible.
- See Lara-Estrada et al. (2024) for more…

Conclusions

- DBN model for multiple flowering events considering feedback between events
- Model captures the flowering dynamic observed in the study region.
- Parameters adjusted for coffee plantations under the climate in the Pacific of C.A.
- The explicit and simple graphical interface in BNs makes the model a practical decision support and learning tool for practitioners and researchers
- Next: Improve and evolve the model into a damage-yield model due to heavy rains during flowering.

References

- Lara-Estrada, L., Sucar, L.E., Rasche, L., 2024. Inferring multiple coffee flowerings in Central America using farmer data in a probabilistic model. Ecological Informatics 79, 102434. https://doi.org/10.1016/j.ecoinf.2023.102434
- Lara-Estrada, L., Haggar, J.P., Ditmar, S., Bruno, R., 2012. Coffee yield variations and their relations to rainfall events in Nicaragua. Presented at the The 24th International Conference on Coffee Science, Association Scientifique Internationale du Café (ASIC), Costa Rica.

For questions or complaints, you can reach me via:
L.LaraEstrada@greenwich.ac.uk

Or find me here